

Perspektiven der modernen Statistik

Autor(en): **Loeffel, Hans**

Objektyp: **Article**

Zeitschrift: **Mitteilungen / Vereinigung Schweizerischer
Versicherungsmathematiker = Bulletin / Association des Actuaire
Suisses = Bulletin / Association of Swiss Actuaries**

Band (Jahr): **72 (1972)**

PDF erstellt am: **27.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-967080>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Perspektiven der modernen Statistik

Von Hans Loeffel, St.Gallen

Meine Ausführungen über Perspektiven moderner Statistik sind im wesentlichen drei Hauptthemen gewidmet:

1. der zunehmenden *Mathematisierung* und Verflechtung der Statistik mit verwandten Wissenschaften, insbesondere der *Entscheidungstheorie*;
2. den fundamentalen Strukturänderungen *technisch-methodischer* Art, bedingt durch die stets wachsende Integration mit der elektronischen Datenverarbeitung;
3. dem Statistikerunterricht an *Mittel- und Hochschulen*.

1. Als *A. N. Kolmogoroff* [1] 1933 die Wahrscheinlichkeitsrechnung als selbständige mathematische Disziplin masstheoretisch begründete, stand auch die mathematische Statistik inmitten einer fruchtbaren Entwicklung, gefördert vor allem von R. A. Fisher einerseits und K. Pearson und J. Neyman andererseits. Besonders hervorzuheben ist R. A. Fisher als Schöpfer der wissenschaftlichen Versuchsplanung. Vergessen wollen wir aber nicht, dass die Wurzeln der modernen Theorie statistischer Inferenz ins 18. Jahrhundert zurückreichen und entscheidend mit dem Namen *Jakob Bernoullis* (I) [2] verbunden sind. Die gedankliche Leistung Bernoullis, niedergelegt in seiner berühmten, erst 1713 veröffentlichten «*Ars conjectandi*», liegt darin, dass er «die Gesamtheit der möglichen statistischen Beobachtungen als ein im Sinne der Wahrscheinlichkeitsrechnung messbares Kollektiv» interpretiert.

Eine nähere Würdigung dieser historischen Leistung findet man bei *Hans Bühlmann* [3] in seinem 1964 erschienenen Artikel «Geburtsstunde der mathematischen Statistik».

In moderner Sicht besteht die Aufgabe des Statistikers ganz allgemein in der *Überwindung* oder Verminderung der *Ungewissheit* im epi-

stemologischen Sinne. Das ist das Inferenzproblem. Die Ungewissheit drückt sich, wenn der Verteilungstyp feststeht, in der Unkenntnis der Parameter aus, die die Verteilung charakterisieren.

Eine bedeutende Wende im statistischen Denken unserer Zeit ist eingetreten durch das Schaffen des genialen Mathematikers *Abraham Wald* (1902–1950). Wald, mit *Oskar Morgenstern* [4] und John von Neumann befreundet, widmete sich neben rein mathematischen Untersuchungen schon früh statistischen und ökonomischen Problemen (z.B. Zeitreihenanalyse, Indexfragen).

Während des Zweiten Weltkrieges gewann in den USA die Qualitätskontrolle im Zusammenhang mit der immensen Kriegsproduktion an Bedeutung. Neue, funktionstüchtige Verfahren wurden gesucht, und Wald leistete dazu einen fundamentalen Beitrag in Form der berühmten *Sequentialanalyse* oder des Sequentialtests [5]. Die 1943 veröffentlichte Studie stand vorerst auf der Geheimliste kriegswichtiger Erfindungen und erregte nach Kriegsende unter den europäischen Mathematikern einiges Aufsehen. Das neuartige Verfahren, auch vom Praktiker leicht zu handhaben, brachte Einsparungen von mehreren Millionen Dollar [4].

Entgegen der Standardpraxis, wo der Stichprobenumfang im voraus feststeht, wird er nach Wald vom Verlauf der Prüfung abhängig gemacht. In jeder Phase des Versuchs wird eine der drei folgenden Entscheidungen getroffen:

- a) Los annehmen,
- b) Los ablehnen,
- c) Fortsetzung der Stichprobe (nächstes Stück ziehen).

Ausgehend von der Sequentialanalyse, entwickelte Wald eine umfassende, mathematisch anspruchsvolle Theorie der sogenannten *statistischen Entscheidungsfunktionen*, die er 1950 in seinem letzten Buch «Statistical Decision Functions» [6] niederlegte. Das statistische Entscheidungsproblem wird als 2-Personen-Nullsummenspiel im Sinne der von Neumannschen Spieltheorie interpretiert. Dabei ist bemerkenswert, dass die Sequentialanalyse einen Spezialfall der Waldschen Entscheidungstheorie darstellt.

Die *klassischen* statistischen Verfahren unterliegen zwei wesentlichen Einschränkungen:

- α) lediglich einmalige Stichprobenentnahme (einstufiges Experiment);
- β) Inferenzprobleme beschränken sich auf Test- und Schätzverfahren ohne Berücksichtigung der ökonomischen Folgen möglicher Entscheidungen.

Wald spannte den Rahmen viel weiter, indem er nicht nur *mehrstufige Entscheidungsprozesse* berücksichtigte, sondern auch die *Versuchsplanung* als Entscheidungsstufe miteinbezog.

Das vereinfachte Modell geht von der Vorstellung aus, dass dem Statistiker oder allgemein dem Entscheidenden gewisse *Aktionen* oder *Letztentscheidungen* a_i zur Verfügung stehen. Andererseits kann die Umwelt oder Realität gewisse *Zustände* ω_j annehmen. Hierbei besteht Unsicherheit darüber, in welchem Zustand sich die Umwelt befindet.

Jedem Zusammentreffen einer Aktion a_i mit einem Zustand ω_j wird ein Verlust $V(a_i, \omega_j)$ zugeordnet.

In der statistischen Entscheidungssituation bedeutet ω_j etwa den Parameter, der eine bestimmte Verteilungsfunktion charakterisiert. *Die Aufgabe des Statistikers besteht nun darin, auf der Basis von beobachteten Zufallsvariablen mittels einer Stichprobe eine bestimmte Aktion a_i zu ergreifen.*

Unter einer Entscheidungsvorschrift oder *Entscheidungsfunktion* versteht man eine Abbildung des Stichprobenraumes in den Raum der Aktionen; mit anderen Worten eine Strategie, die jeder möglichen Realisation der Zufallsvariablen eine bestimmte Aktion oder Letztentscheidung zuordnet.

Jeder Entscheidungsfunktion kann bei gegebenem Zustand ω_j der *mittlere Verlust* oder das sogenannte *Risiko* zugeordnet werden. Ein *Entscheidungskriterium* liefert uns dann die in einem gewissen Sinne *optimale* Strategie. Falls die Realität (Parameter der unbekanntem Verteilungsfunktion) als rationaler Gegenspieler auftritt, kann die optimale Entscheidungsfunktion nach dem *Minimax-Kriterium* ermittelt werden durch Minimierung des maximalen Risikos. Kann über die Häufigkeit des Auftretens der verschiedenen Zustände eine Aussage gemacht werden, mit anderen Worten, kann über dem Zustandsraum Ω eine sogenannte *A-priori-Wahrscheinlichkeitsverteilung* als bekannt vorausgesetzt werden, so erfolgt die Anwendung des sogenannten *Bayesschen Kriteriums*. Danach gilt jene Strategie als die beste, für die das erwartete Risiko (bezüglich der Verteilung über Ω) möglichst klein ist.

Obschon die entscheidungsorientierte Statistik den wissenschaftlichen Horizont erweiterte und die Eingliederung in die ökonomische Theorie ermöglichte, waren die bereits vorhandenen Kontroversen epistemologischer Art nicht verstummt.

Offene Fragen bestehen vor allem noch in folgenden Bereichen:

- A. Sind sogenannte *subjektive Wahrscheinlichkeiten* zulässig?
- B. Sind die Informationen zur Festlegung der *Verlustfunktion* $V(a_i, \omega_j)$ hinreichend genau?
- C. Welches *Entscheidungskriterium* soll gewählt werden?
- D. Ist das Entscheidungsmodell *operationell* geeignet?

A. Die Verwendung des Bayesschen Kriteriums setzt die Kenntnis einer A-priori-Verteilung über dem Zustandsraum Ω voraus, die oft objektiv, d. h. häufigkeitsinterpretiert, nicht gegeben ist. Massgebende Wahrscheinlichkeitstheoretiker und Statistiker wie *L. J. Savage* [7], *de Finetti*, *R. Schlaifer* u. a. m. befürworteten deshalb die Verwendung von subjektiven Wahrscheinlichkeiten, die als Mass für die Glaubwürdigkeit zu interpretieren sind, mit denen die Ereignisse eintreffen. Die Wahrscheinlichkeiten lassen sich in gewisser Hinsicht definieren durch das potentielle Verhalten einer Person in einer Wettsituation.

C. Je nach dem verwendeten Entscheidungskriterium kann die optimale Strategie verschieden ausfallen. Wie soll in einer konkreten Situation das angemessene Entscheidungskriterium gefunden werden?

D. *Günter Menges* [8] hat vor etwa fünf Jahren die Frage der operationellen Eignung von Entscheidungsmodellen aufgeworfen. Ein solches heisst *operabel*, wenn es zu seiner konkreten Anwendung nur Informationen oder Daten erfordert, die auch tatsächlich gewonnen werden können.

Auch das Mehrdeutigkeits- und Vorentscheidungsproblem sowie die Dynamisierung spielen in diesem Zusammenhang eine Rolle.

Die entscheidungstheoretisch orientierte Statistik ist nicht unumstritten. *R. A. Fisher* z. B. hat sich in seinen letzten Lebensjahren gegen jegliche Anwendung von statistischen Entscheidungskalkülen gewandt.

Während die klassische Inferenztheorie *Methoden* für die Überwindung der Unsicherheit liefert, gibt uns die Entscheidungstheorie *Richtlinien* für das rationale Verhalten in ungewissen Situationen. Ein

einseitiger Standpunkt kann aber sicher nicht *allen* statistischen Problemen gerecht werden.

Wir brauchen beides:

Stets verfeinerte praxis- und problemorientierte *Methoden* einerseits, aber auch *grundsätzliche* Untersuchungen über die Aussagefähigkeit statistischer Theorien andererseits.

Neben der Entwicklungstendenz in Richtung der Entscheidungstheorie ist ein wachsendes Interesse an sogenannten *verteilungsfreien* Verfahren zu erkennen, die besonders in den Wirtschafts- und Sozialwissenschaften fruchtbare Anwendung finden. Seit ungefähr 1940 werden immer häufiger Arbeiten publiziert, welche nicht mehr von der üblichen Voraussetzung der Normalität ausgehen. Die ersten Arbeiten gehen auf *S. S. Wilks* zurück (1938–39) [9]. Die Grundidee liegt in der *Randomisierung der experimentellen Daten*, die es erlaubt, die Wahrscheinlichkeitsrechnung (und damit Aussagen über die Normalverteilung) zur Interpretation der Daten heranzuziehen, ohne annehmen zu müssen, dass dieselben aus normal verteilten Grundgesamtheiten stammen. Im Heft 2 des 70. Bandes der Mitteilungen findet sich eine Arbeit von Dr. Berliner über die Anwendbarkeit von *Anordnungstests* im Versicherungswesen. Diese können eine wichtige *Kontrollfunktion* ausüben, indem sie angeben, wann ein bestehendes Prämiensystem aus dem Gleichgewicht gerät.

2. Im folgenden wollen wir uns kurz der *technischen* und *anwendungsorientierten Seite* zuwenden.

Der bekannte Kybernetiker Prof. *Karl Steinbuch* [10] schreibt in seinem Buch «Falsch programmiert»:

«Vermutlich ist der Computer dasjenige technische Phänomen, welches das Denken unserer Zeit am stärksten verändert.»

In der Tat glaube ich, dass im Bereiche der angewandten Statistik der Grosscomputer von heute und jener der nahen Zukunft Möglichkeiten der Datenerfassung, Speicherung und Auswertung eröffnet, die noch kaum zu überblicken sind. Insbesondere hat die *beschreibende Statistik* in letzter Zeit an Bedeutung gewonnen, nicht zuletzt im Sektor der internationalen Statistik, die einen *weltweiten Informationsaustausch* gewährleisten soll.

Vermutlich wird in gewissen Bereichen (amtliche und Betriebsstatistik) die statistische Auswertungspraxis stark modifiziert werden in

Anbetracht der Tatsache, dass in Grossspeichern eine Unmenge von Daten festgehalten werden kann, um jederzeit wieder individuell oder kollektiv darüber zu verfügen. Diese sogenannten *Datenbanken* werden in Zukunft eine wichtige, vielleicht auch gefährliche Entscheidungsgrundlage bilden.

Eine besonders erfreuliche Perspektive eröffnet sich in der Tatsache, dass der Computer immer mehr als Fertiglieferant zahlreicher statistischer *Techniken* unentbehrliche Dienste leistet. Die Zeiten sind wohl endgültig vorbei, in denen der Statistiker in mühevoller, monotoner Arbeit die Verfahren durchexerzieren musste.

Gegenwärtig sind *zwei* bedeutsame Entwicklungen zu beobachten. Einmal gelangen immer anspruchsvollere Tischcomputer (von der Grössenordnung einer Schreibmaschine) auf den Markt, die eine Anzahl statistischer Routinen fest programmiert enthalten. Ausserdem sind sie oft zusätzlich von aussen noch programmierbar. Auf der andern Seite gibt die *Datenfernverarbeitung* die Möglichkeit (z.B. mit Time-sharing) jederzeit die gewaltige Kapazität und das reiche Angebot der Software eines angeschlossenen Grosscomputers auszunützen. Direkt vom Arbeitsplatz aus, wo ein sogenanntes *Terminal* (Eingabe-Ausgabe-Gerät) untergebracht ist, verfügt der Interessent in direktem Zugriff über eine grosse Auswahl statistischer *Techniken* (z.B. Regressionsanalyse, klassische Test- und Schätzverfahren).

Wie Prof. Bauknecht an der diesjährigen Jahresversammlung der Gesellschaft für Volkswirtschaft und Statistik ausgeführt hat, besteht ein dringendes Bedürfnis in einem weiteren Ausbau einer sogenannten *Methodenbank*, die ein integraler Bestandteil eines wirkungsvollen Informationssystem ist.

Trotz allem berechtigten Optimismus ist eine Warnung angebracht. Es besteht die Verlockung, allzu rasch in kritikloser und naiv-unbesorgter Art statistische Verfahren mit dem Computer «durchzuspielen», nicht selten von der trügerischen Hoffnung getragen, wissenschaftliche Arbeit geleistet zu haben.

3. Zum Schluss noch einige Bemerkungen zu Fragen des Unterrichts in Statistik.

Die moderne Statistik ist eine interdisziplinäre Methodenlehre der Überwindung von Ungewissheit durch induktive Schlüsse, die ihre Basis in empirischen Betrachtungen haben. Fast jeder Wissenszweig,

der die Realität im weitesten Sinne modellhaft zu erklären versucht, wird mit der Statistik konfrontiert. Daraus ziehen wir die Konsequenz, dass die Frage der *Grundausbildung* auf diesem Gebiet von grösster Bedeutung ist, und zwar sowohl auf Mittel- wie auch auf Hochschulstufe.

An den meisten öffentlichen *Mittelschulen* wird heutzutage gemäss den mehr oder weniger verbindlichen Weisungen der Eidgenössischen Maturitätskommission eine *elementare Einführung in die Wahrscheinlichkeitsrechnung* vermittelt. Im Sinne eines modernen, d. h. zukunfts- und praxisorientierten Unterrichts scheint mir ein angemessener Ausbau der Wahrscheinlichkeitsrechnung und der beschreibenden Statistik bereits auf der Mittelschulstufe angebracht. Das neue Wirtschaftsgymnasium (nach den Leitideen von Prof. Dubs), das an einigen Schulen schon verwirklicht ist, wird auch den künftigen Wirtschaftswissenschaftlern jenes mathematische Grundwissen vermitteln, das man ihnen viel zu lange vorenthalten hat.

Eine allzu breite Auswalzung gewisser Spezialgebiete der Funktionenlehre und analytischen Geometrie sollte vermieden werden im Interesse einer elementaren Einführung in das *stochastische Denken*. Diese Aufgabe erfordert vom Lehrer viel didaktisches Geschick, weil weniger die routinemässige Technik als vielmehr die intuitive Durchdringung und wechselseitige Interpretation im Vordergrund stehen.

Die interdisziplinäre Bedeutung der modernen Statistik macht sich auch in zunehmendem Masse auf der *Hochschule* bemerkbar. In dieser Frage möchte ich mich auf jene Hochschulen beschränken, die Wirtschafts- und Sozialwissenschaften lehren.

Noch vor ungefähr zehn Jahren wurde an solchen Lehrstätten die *mathematische* Statistik kaum, oder dann nur in schwachen Dosen, verabreicht. Dies wohl infolge der fehlenden mathematischen Grundbildung und der mehr oder weniger unbewusst vorhandenen Aversion gegen alle Quantifizierung und Formalisierung.

Diese Aversion wird sicher in dem Masse abnehmen, wie die Fortschritte im Operations Research, in der Ökonometrie und der elektronischen Datenverarbeitung zunehmen. Auch die Statistik wird im Zuge dieser Entwicklung in der zweckgerichteten Forschung einen immer breiteren Platz einnehmen als ein Hilfsmittel der rationalen Entscheidungsbildung.

Aber auch abgesehen von jeglichen Nützlichkeitsüberlegungen, nimmt die Statistik und das damit verbundene stochastische Denken eine Schlüsselstellung ein in der *grundsätzlichen* Erfassung und Einstufung der modernen *empirischen* Wirtschafts- und Sozialwissenschaften.

Literaturverzeichnis

- [1] *A. N. Kolmogoroff*: Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin 1933.
- [2] *Jakob Bernoulli*: Ars conjectandi, Basel 1713.
- [3] *Hans Bühlmann*: Die «Geburtsstunde» der mathematischen Statistik, in: Vierteljahresschrift der Naturforschenden Gesellschaft Zürich, Jahrgang 109, Heft 3, 1964.
- [4] *Oskar Morgenstern*: Spieltheorie und Wirtschaftswissenschaft, R. Oldenbourg-Verlag, Wien 1963 (enthält im Anhang einen kurzen Lebenslauf von Abraham Wald).
- [5] *A. Wald*: Sequential Analysis, Wiley-Verlag, New York 1947.
- [6] Statistical Decision Functions, Wiley-Verlag, New York 1950.
- [7] *L. J. Savage*: The Foundations of Statistical Inference, Methuen's Monographs, London 1962.
- [8] *G. Menges/H. Diehl*: Über die operationelle Eignung von Entscheidungsmodellen, in: Statistische Hefte, Heft 1/2, 1966, S. 30–41.
- [9] *S. S. Wilks*: Mathematical Statistics, New York/London 1962.
- [10] *Karl Steinbuch*: Falsch programmiert, Deutscher Taschenbuch Verlag, München 1969.
- [11] *Hans Riedwyl*: Statistik im Wandel der Zeit, in: Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, 70. Band, Heft 2.

Zusammenfassung

Es wird vorerst gezeigt, wie die moderne Statistik im Laufe der letzten Jahrzehnte für fast alle empirischen Wissenschaften an Bedeutung zugenommen hat. Als allgemeine Methodenlehre der Überwindung von Ungewissheit steht sie in enger Tuchfühlung mit der mathematischen Entscheidungstheorie.

Es folgen Hinweise auf die zahlreichen strukturellen Veränderungen, die der moderne Computer in der beschreibenden und operationellen Statistik hervorgerufen hat.

Abschliessend setzt sich der Autor für einen zeitgemässen Ausbau des Statistikerunterrichts an Mittel- und Hochschulen ein. Dies nicht zuletzt deshalb, weil das mit der Statistik verbundene stochastische Denken auch in wissenschaftstheoretischer Sicht eine Schlüsselstellung einnimmt.

Summary

The author begins by stressing the fact that in recent times Modern Statistics has become more and more important for nearly all the empirical sciences.

As the general science of decision-making in the face of uncertainty it is closely related to the theory of mathematical decision.

The next paragraph deals with the manifold structural changes caused by the modern computer.

The author finally pleads for an improvement in the methods of teaching Statistics at high schools and universities, as there is no denying that random structures, which are bound up with Statistics, are of great influence in the field of the theory of science too.

Résumé

Tout d'abord on démontre l'importance croissante de la statistique moderne dans le domaine des sciences empiriques au cours des dernières décennies. En tant que méthodologie générale, pour éliminer les incertitudes, elle est en rapport étroit avec la théorie mathématique des décisions.

Par la suite on donne des références aux nombreuses modifications structurales dues à l'utilisation de l'ordinateur dans la statistique descriptive et opérationnelle.

Finalement l'auteur intervient en faveur d'une expansion conforme aux exigences actuelles de l'enseignement de la statistique dans les écoles du degré supérieur et universitaire. Ceci en raison du fait que les structures aléatoires liées à la statistique occupent actuellement une position clef dans l'élaboration des théories scientifiques.

Riassunto

Si dimostra dapprima come negli ultimi decenni l'importanza della statistica sia cresciuta per quasi tutte le scienze empiriche. Come teoria generale per superare l'incertezza la statistica è in stretta relazione con la teoria matematica delle decisioni.

Seguono degli accenni ai numerosi cambiamenti strutturali provocati dal calcolatore elettronico nella statistica descrittiva e operativa.

Infine l'autore si esprime a favore di uno sviluppo dell'insegnamento della statistica nelle scuole medie e superiori, che corrisponde alle esigenze moderne. Questo anche per il fatto che le strutture probabilistiche, legate alla statistica, occupano una posizione chiave nell'elaborazione di teorie scientifiche.

